American Journal of Medical Genetics (Neuropsychiatric Genetics) 67:495–498 (1996)

Brief Research Communication

Analysis of Polyglutamine-Coding Repeats in the TATA-Binding Protein in Different Human Populations and in Patients With Schizophrenia and Bipolar Affective Disorder

David C. Rubinsztein, Jayne Leggo, Timothy J. Crow, Lynn E. DeLisi, Cathy Walsh, Sanjeev Jain, and Eugene S. Paykel

East Anglian Medical Genetics Service Molecular Genetics Laboratory, Addenbrooke's National Health Service Trust, Cambridge (D.C.R., J.L.); Department of Psychiatry, Cambridge University, Addenbrooke's Hospital, Cambridge (S.J., C.W., E.S.P.); Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford (T.J.C.), United Kingdom; National Institute of Mental Health and Neurosciences, Bangalore (S.J.), India; Department of Psychiatry and Behavioral Science, State University of New York at Stony Brook, Stony Brook, New York (L.E.D.)

A new class of disease (including Huntington disease, Kennedy disease, and spinocerebellar ataxias types 1 and 3) results from abnormal expansions of CAG trinucleotides in the coding regions of genes. In all of these diseases the CAG repeats are thought to be translated into polyglutamine tracts. There is accumulating evidence arguing for CAG trinucleotide expansions as one of the causative disease mutations in schizophrenia and bipolar affective disorder. We and others believe that the TATA-binding protein (TBP) is an important candidate to investigate in these diseases as it contains a highly polymorphic stretch of glutamine codons, which are close to the threshold length where the polyglutamine tracts start to be associated with disease. Thus, we examined the lengths of this polyglutamine repeat in normal unrelated East Anglians, South African Blacks, sub-Saharan Africans mainly from Nigeria, and Asian Indians. We also examined 43 bipolar affective disorder patients and 65 schizophrenic patients. The range of polyglutamine tract-lengths that we found in humans was from 26–42 codons. No patients with bipolar affective disorder and schizophrenia had abnormal expansions at this locus.

© 1996 Wiley-Liss, Inc.

KEY WORDS: schizophrenia, bipolar affective disorder, trinucleotide repeats, TATA binding protein

There has been considerable interest recently in a new class of disease resulting from abnormal expansions of CAG trinucleotides in the coding regions of genes. These diseases, which include Huntington disease, spinocerebellar ataxias types 1 and 3, Kennedy disease, and dentatorubral-pallidoluysian atrophy, share a number of features [reviewed by Ross, 1995]. In all these diseases the CAG repeats are thought to be translated into polyglutamine tracts. While these repeats are polymorphic on normal chromosomes, they are generally stable when transmitted from one generation to the next. However, CAG repeats associated with disease chromosomes have high mutation rates and have a bias towards increasing repeat numbers when passed from one generation to the next. Since age of onset of these diseases correlates inversely with number of CAG repeats on disease chromosomes, these diseases tend to manifest at earlier ages in successive generations, a phenomenon known as anticipation.

There is accumulating evidence arguing for CAG trinucleotide expansions as one of the causative disease mutations in schizophrenia and bipolar affective disorder. These diseases have been reported to show clinical features of anticipation in pedigrees [e.g., Basset and Honer, 1994; McInnis et al., 1993], and experiments using the Repeat Expansion Detection (RED) technique have suggested that patients with bipolar affective dis-

Received for publication January 16, 1996; revision received February 29, 1996.

Address reprint requests to David C. Rubinsztein, East Anglian Medical Genetics Service Molecular Genetics Laboratory, Box 158, Addenbrooke's National Health Service Trust, Hills Road, Cambridge CB2 2QQ, UK.

© 1996 Wiley-Liss, Inc.

order and schizophrenia show larger CAG repeats in their genome than controls [O'Donovan et al., 1995; Morris et al., 1995; Lindblad et al., 1995].

We and others [Imbert et al., 1994; Rosen et al., 1995] believe that the TATA-binding protein (TBP) is a compelling candidate for this class of diseases for the following reasons. First, the TATA-binding protein is a critical factor in transcriptional initiation and is ubiquitously expressed [reviewed by Gostout et al., 1993; Imbert et al., 1994]. Second, it contains a stretch of uninterrupted glutamine codons which show similar degrees of polymorphism in the normal population to the five known diseases. (Although the codons all code for glutamine and are thought to be translated, they are not all CAG.) Third, most normal human TBP alleles have more than 35 glutamine codons and are thus close to the threshold length where the polyglutamine tracts start to be associated with disease. For instance, Huntington disease alleles have been described with 38 uninterrupted glutamines (equivalent to 36 CAG repeats followed by CAACAG) Rubinsztein et al. [in press]. Fourth, Trottier et al. [1995] recently showed that a monoclonal antibody recognizing peptides overlapping the polyglutamine tract in TBP selectively recognizes only the expanded pathogenic gene products in cells from patients with Huntington disease and spinocerebellar ataxias 1 and 3.

We were interested in exploring the size ranges of TBP repeats in humans, in looking for alleles with long polyglutamine tracts, and in investigating this locus for abnormal expansions in patients with schizophrenia and bipolar affective disorder. Thus, we examined the polyglutamine repeat length in unrelated East Anglian cystic fibrosis patients who were not homozygous for the ΔF508 mutation (49 individuals), normal South African Blacks (48), sub-Saharan Africans mainly from Nigeria referred for diagnosis of sickle-cell anemia (47), and Asian Indians referred for diagnosis of hemoglobinopathies (45) (details of populations can be found in Rubinsztein et al. [1994]). A number of normal populations were examined in order to sample a wider range of alleles. We also examined 43 bipolar affective disorder patients and 65 schizophrenic patients.

Unrelated schizophrenic patients were probands from families with at least 2 sibs with either schizophrenia or schizoaffective disorder identified in North West London and its surroundings, Northern Italy, and several states in the USA. All subjects were evaluated by a trained clinician using a modified Schedule for Affective Disorders and Schizophrenia (SADS) structured format [Endicott and Spitzer, 1978]. Records from previous hospital admissions and psychiatric treatment diagnoses were obtained, and further information was collected from reliable family members. Diagnoses were made based on these sources using DSM-III-R criteria DSM III R American Psychiatric Association [1989]. Unrelated bipolar affective disorder patients were recruited from inpatient and outpatient clinics in East Anglia. The sample met Research Diagnostic Criteria for bipolar affective disorder, type 1 [Spitzer et al., 1978]. Lifetime psychopathology was assessed using the SADS-L (Lifetime version) [Endicott and Spitzer, 1978]. Ethical approval was granted by the Oxfordshire Psychiatric Research Ethics Committee and the Ethics Committees at Stony Brook University and Addenbrooke's Hospital for genetic analysis in patients with schizophrenia and bipolar affective disorder.

If TBP caused schizophrenia or bipolar affective disorder in a way that was analogous to that observed in all other trinucleotide repeat diseases, one would expect distinct normal and disease allele distributions. We previously modelled the probabilities of detecting a significant excess of large expansions in scenarios where 1/50 individuals in control groups were assumed to carry schizophrenia or bipolar affective disorder susceptibility genes, and a varying proportion of cases of either disease was associated with a trinucleotide repeat expansion [Jain et al., 1996]. Simulation experiments demonstrated that this approach for analysis of trinucleotide-containing genes has the power to exclude the candidate as being responsible for >33% of cases with the disease, even if only 20 cases and 20 controls are examined [Jain et al., 1996]. In this case we examined many more individuals.

We analyzed polyglutamine lengths in TBP using primers TBP1 (5′-CTG TCT ATT TTG GAA GAG CAA CAA AGG) and TBP2 (5′-CTG CTG GGA CGT TGA CTG CTG AAC G) [Polymeropoulos et al., 1991]. PCR reactions were performed in 12.5 μl, containing 30 ng of each primer; 1 × PCR buffer (50 mM KCl, 10 mM Tris, pH 8, 0.01% gelatin, 1.5 mM $MgCl_2$); 10% DMSO; 200 μM dATP, dCTP, and dTTP; 50 μM dGTP; 150 μM 7-deaza dGTP; about 0.5 μg DNA; and 0.3 ng of 5′ end-labelled TBP1. Tubes were denatured at 95°C for 5 min *Taq* polymerase was added, and the tubes were cycled 35 times through 95°C (30 sec), 62°C (30 sec), and 72°C (60 sec). Sizes of alleles were determined by comparison to a known sequence ladder and confirmed by direct sequencing of selected PCR products with the *f*mol (Promega, Madison, WI) kit using primers TBP1 and TBP2.

The range of polyglutamine tract-lengths in humans was from 26–42 codons (Table I). Sequencing of a number of alleles confirmed that the length polymorphism was due to a variable number of consecutive glutamine codons (not all coded for by CAG).

Allele lengths in the bipolar and schizophrenic patients were confined to the normal range. The largest alleles in these patients had 42 glutamines. Such an allele with 42 consecutive glutamines was found in one of our Asian Indian controls, and once each in two previous studies [Gostout et al., 1993; Imbert et al., 1994]. A 42-repeat allele was shown to segregate in seven meioses in a Centre d'Etude du Polymorphisme Humain (CEPH) family, and therefore has been found in a fairly large number of apparently normal individuals [Imbert et al., 1994]. It is unlikely that abnormally long alleles in TBP are found in the majority of patients with schizophrenia and bipolar affective disorder. The possibility that we missed large alleles refractory to PCR amplification is small, as gels were exposed to X-ray film longer than necessary for the detection of normal alleles, and most of the patients were heterozygotes (see Table I).

TABLE I. Percentage of Chromosomes With Different TBP Polyglutamine Lengths in Different Control Human Populations and in Schizophrenic and Bipolar Affective Disorder Patients*

No. of glutamines	East Anglians (n = 49)	South African Blacks (n = 48)	Sub-Saharan Africans (n = 47)	Asian Indians (n = 45)	Bipolar (n = 43)	Schizophrenic (n = 65)
24						
25						
26					1	
27				2		
28						
29						
30						1
31				1	1	
32	3					3
33	2	10	19			
34		22	18			2
35	8	34	24	6	2	5
36	14	18	11	19	12	13
37	23	8	12	22	29	19
38	43	5	13	30	42	43
39	7	3	3	19	7	12
40					5	2
41						
42				1	1	1
% of heterozygotes	84	71	85	78	84	78

* Total number of chromosomes and percentage of heterozygotes in each population is indicated.

Because Huntington disease, spinocerebellar ataxia types 1 and 3, and Kennedy disease share a common mutation (which seems to act as a "gain of function" at the protein level) and certain clinicopathological features [Ross, 1995], a number of models have proposed a common pathway where the polyglutamine tracts cause disease when they exceed a critical length. Green [1993] suggested that long polyglutamine tracts could be substrates for transglutaminases, which can cross-link these proteins to other similar molecules which contain glutamines, or which can serve as lysyl donors. Perutz et al. [1994] showed that polyglutamines can act as polar zippers and suggested that when these tracts reach pathogenic lengths they react nonspecifically with other glutamine-containing proteins, which ultimately precipitate. Cha and Dure [1994] hypothesized that mutant proteins would be processed to give rise to polyglutamine-containing peptides, which may ultimately act as "toxic" degradation products.

Forty-two consecutive glutamines would unequivocally represent a disease allele for Huntington disease, spinocerebellar ataxia type 1, and Kennedy disease [Rubinsztein et al., in press; Doyu et al., 1993; Genis et al., 1995]. Accordingly, the simplest models whereby expanded polyglutamines cause disease will have to be altered to accommodate the observation of apparently normal individuals with 42 TBP repeats. It is possible that the context of polyglutamines in the tertiary protein structure may modify the length dependency of their "toxicity" or the way that they bind to interacting proteins. Alternatively, the rates of synthesis or the degradation of different polyglutamine-containing proteins in different tissues may in some way affect their pathogenic pathways (and may account for the distinct patterns of neurodegeneration in different polyglutamine diseases).

In conclusion, our data suggest that expansions of the polyglutamine tracts in TBP are not responsible for the majority of cases of schizophrenia and bipolar affective disorder. Recently, this gene was mapped to 6q27 [Imbert et al., 1994; Rosen et al., 1995]. We believe that it still remains an important candidate for other diseases showing features of anticipation, which have not yet been conclusively mapped, or which map to this region.

ACKNOWLEDGMENTS

We thank Drs. A. Vita, C. Morganti, J. Old, and R. Ramesar for DNA samples. S.J. was supported by a Commonwealth Fellowship. Part of this work was supported by grants from the Huntington's Disease Association, UK (to D.C.R.), from the Medical Research Council (to T.J.C.), and from the NIMH R01-844245 (to L.E.D.L.).

REFERENCES

Basset AS, Honer WG (1994): Evidence for anticipation in schizophrenia. Am J Hum Genet 54:864–870.

Cha JH, Dure LS IV (1994): Trinucleotide repeats in neurologic diseases: An hypothesis concerning the pathogenesis of Huntington's disease, Kennedy's disease, and spinocerebellar ataxia type I. Life Sci 54:1459–1464.

Doyu M, Sobue G, Mitysuma T, Uchida M, Iwase T, Takahashi A (1993): Very late onset recessive bulbospinal neuronopathy: Mild clinical features and a mild increase in the size of the tandem CAG repeat in the androgen receptor gene. J Neurol Neurosurg Psychiatry 56:832.

DSM III R American Psychiatric Association (1989): "Diagnostic and Statistical Manual of Mental Disorders" 3rd ed. revised.

Endicott J, Spitzer RL (1978): A diagnostic interview: The Schedule for Affective Disorders and Schizophrenia. Arch Gen Psychiatry 35:837–862.

Genis D, Matilla T, Volpini V, Rosell J, Davalos A, Ferrer I, Molins A, Estivill X (1995): Clinical, neuropathologic, and genetic studies of a large spinocerebellar ataxia type 1 (SCA 1) kindred: (CAG)n expansion and early premonitory signs and symptoms. Neurology 45:24–30.

Gostout B, Liu Q, Sommer SS (1993): "Cryptic" repeating of purines and pyrimidines (cRRY(i)) are frequent and polymorphic: analysis of coding cRRY(i) in the proopiomelanocortin (POMC) and TATA-binding protein (TBP) genes. Am J Hum Genet 52:1182–1190.

Green H (1993): Human genetic diseases due to codon reiteration: Relationship to an evolutionary mechanism. Cell 74:955–956.

Imbert G, Trottier Y, Beckmann J, Mandel JL (1994): The gene for the TATA binding protein (TBP) that contains a highly polymorphic protein coding CAG repeat maps to 6q27. Genomics 21:667–668.

Jain S, Leggo J, DeLisi LE, Crow TJ, Margolis RL, Li S-H, Goodburn S, Walsh C, Paykel E, Ferguson-Smith MA, Ross CA, Rubinsztein DC (1996): Analysis of thirteen trinucleotide repeat loci as candidate genes for schizophrenia and bipolar affective disorder. Am J Med Genet 67:139–146.

Lindblad K, Nylander PO, Debruyn A, Sourey D, Zander C, Engstrom C, Holmgren G, Hudson T, Chotai J, Mendlewicz J, Van Broekhoven C, Schalling M, Adolfsson R (1995): Detection of expanded CAG repeats in bipolar affective disorder using the repeat expansion detection (RED) method. Neurobiol Dis 2:55–62.

McInnis MG, McMahon FJ, Chase GA, Simpson SG, Ross CA, DePaulo JR (1993): Anticipation in bipolar affective disorder. Am J Hum Genet 53:385–390.

Morris AG, Gaitonde E, McKenna PJ, Hunt DM (1995): CAG repeat expansions and schizophrenia: Association with disease in females and with early age-at-onset. Hum Mol Genet 4:1957–1961.

O'Donovan MC, Guy C, Craddock N, Murphy KC, Cardno AG, Jones LA, Owen MJ, McGuffin P (1995): Expanded CAG repeats in schizophrenia and bipolar disorder. Nat Genet 10:380–381.

Perutz MF, Johnson T, Suzuki M, Finch JT (1994): Glutamine repeats as polar zippers: Their possible role in inherited neurodegenerative disease. Proc Natl Acad Sci USA 91:5355–5358.

Polymeropoulos MH, Rath DS, Xiao H, Merril CR (1991): Trinucleotide repeat polymorphism at the human transcription factor IID gene. Nucleic Acids Res 19:4307.

Rosen DR, Trofatter JA, Brown RH Jr (1995): Mapping of the human TATA-binding protein gene (TBP) to chromosome 6qter. Cytogenet Cell Genet 69: 279–280.

Ross CA (1995): When more is less: pathogenesis of glutamine repeat neurodegenerative diseases. Neuron 15:493–496.

Rubinsztein DC, Amos W, Leggo J, Goodburn S, Ramesar R, Old J, Bontrop R, McMahon R, Barton DE, Ferguson-Smith MA (1994): Mutational bias provides a model for the evolution of Huntington's disease and predicts a general increase in disease prevalence. Nat Genet 7:525–530.

Rubinsztein DC, Leggo J, Coles R, et al. (in press): Phenotypical characterization of individuals with 30–40 CAG repeats in the Huntington disease (HD) gene reveals HD cases with 36 repeats and apparently normal elderly individuals with 36–39 repeats. Am J Hum Genet.

Spitzer RL, Endicott J, Robins E (1978): Research Diagnostic Criteria: Rationale and reliability. Arch Gen Pcychiatry 35:773–782.

Trottier Y, Lutz Y, Stevanin G, Imbert G, Devys D, Cancel G, Saudou F, Weber C, David G, Tora L, Agid Y, Brice A, Mandel J-L (1995): Polyglutamine expansion as a pathological epitope in Huntington's disease and four dominant cerebellar ataxia. Nature 378: 403–406.